Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

What drives failure to maximize payoffs in the lab ?
A test of the inequality aversion hypothesis

Nicolas JACQUEMET, Adam ZYLBERSZTEJN

2011.36

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/bandeau-haut/documents-de-travail/
ISSN : 1955-611X

What drives failure to maximize payoffs in the lab? A test of the inequality aversion hypothesis*

Nicolas Jacquemet† Adam Zylbersztejn*

June 2011

*We wish to thank Nick Feltovitch and Stéphane Luchini for their valuable comments. We acknowledge financial support from the University Paris 1 Panthèon-Sorbonne and the Paris School of Economics.

†Paris School of Economics and University Paris I Panthéon–Sorbonne. Centre d'Economie de la Sorbonne, 106 Bd. de l'Hopital, 75013 Paris, France. Nicolas.Jacquemet@univ-paris1.fr ; adam.zylbersztejn@malix.univ-paris1.fr.

Abstract

In experiments based on the Beard and Beil (1994) game, second movers very often fail to select the decision that maximizes both players payoff. This note reports on a new experimental treatment, in which we neutralize the potential effect of inequality aversion on the likelihood of this behavior. We show this behavior is robust to this change, even after allowing for repetition-based learning.

Keywords: Coordination Failure, Laboratory experiments, Aversion to inequality. **JEL Classification:** C72, D83.

Résumé

Dans les expériences en laboratoire fondées sur le jeu de Beard et Beil (1994), les joueurs chargés de décider en second échouent très souvent à prendre la décision qui maximise simultanément les gains des deux joueurs en présence. Ce court article présente les résultats d'une expérience dont le protocole neutralise les effets potentiels de l'aversion à l'inégalité. Les comportements observés sont tout à fait robustes à ce changement dans l'environnement, y compris après un certain nombre de répétitions du jeu statique.

Keywords: Coordination Failure, Laboratory experiments, Aversion to inequality.

1 Introduction

Causal inference from laboratory experiments relies on the crucial property that subjects' behavior is driven by financial incentives, in such a way that *ceteris paribus* a subject always prefers more money to less. This requirement is key to guarantee that the environment the experimenter decided to implement is actually the one in which decisions are made. The seminal study by Beard and Beil (1994) reports on experimental results which tend to challenge this view. They rely on a sequential two-players one-shot game originally introduced by Rosenthal (1981) in which the first mover either decides alone on the final issue of the game or relies on the second mover, in which case the second mover only has to decide whether to maximize both players' payoff or not. But since the first mover looses a lot should the second mover fail to maximize payoffs, a very likely issue is the Pareto dominated one in which the first mover decides alone. Two features emerge from the existing experimental implementations of this game (summarized in Section 2 below). First, the Pareto-suboptimal outcome arises very often; second, it is largely because an important share of the subjects playing as second movers indeed fail to maximize payoffs when they have the opportunity to do so.

This paper investigates one possible explanation for this striking behavior, namely aversion to inequality. In contrast to previous studies of this game, we neutralize the potential effect of such kind of preferences using a perfectly symmetric payoff structure. We show that this change in the game has very few consequences on subjects' behavior. We conclude the paper with a discussion of open avenues to better understanding why subjects to (these) laboratory experiments may fail to maximize payoff.

2 Related literature

Table 1 provides an overview of experimental implementations of the game we study. There are only three different outcomes in the game, detailed in the left-hand side of the Table: either the first mover (henceforth *player A*) chooses to decide alone by picking up L, or he relies on the second mover's (henceforth *player B's*) decision by choosing R. In this case, both players' payoffs are higher if r is chosen rather than l. The right-hand side of the Table summarizes the share of each outcome among all observed decisions, as well as the frequency of action r conditional on reliance from player A.

The main focus of Beard and Beil (1994) is to test the conjecture made by Rosenthal that subjects may be reluctant to rely on other's ability to maximize payoffs – hence challenging subgame perfectness.[1] Their early experimental evidence supports the Rosenthal conjecture: while the share of payoff maximizing subjects is high in the sub-population of player Bs who are relied

[1] It is so because in this particular game such behavior amounts to use weakly dominated strategies. See Jacquemet and Zylbersztejn (2010) for a more detailed analysis of the theoretical properties of the game.

Table 1: Summary of experimental evidence on Rosenthal's game

Experiment	Payoff			Observed outcomes				Nb.
	(L)	(R,l)	(R,r)	L	(R,r)	(R,l)	$Pr(r\|R)$	obs.
Beard & Beil (1994), Tr.1	(9.75; 3)	(3; 4.75)	(10; 5)	66%	29%	6%	83%	35
Beard & Beil (1994), Tr.3	(7.00; 3)	(3; 4.75)	(10; 5)	20%	80%	0%	100%	25
Beard & Beil (1994), Tr.4	(9.75; 3)	(3; 3.00)	(10; 5)	47%	53%	0%	100%	32
Beard et al. (2001), Tr.1	(1450; 450)	(450; 700)	(1500; 750)	79%	18%	3%	83%	34
Beard et al. (2001), Tr.2	(1050; 450)	(450; 700)	(1500; 750)	50%	18%	32%	64%	28
Goeree & Holt (2001), Tr.1	(80; 50)	(20; 10)	(90; 70)	16%	84%	0%	100%	25
Goeree & Holt (2001), Tr.2	(80; 50)	(20; 68)	(90; 70)	52%	36%	12%	75%	25
Baseline, round 1	(9.75; 3)	(3; 4.75)	(10; 5)	77%	23%	0%	100%	30
Baseline, rounds 2-10	(9.75; 3)	(3; 4.75)	(10; 5)	48%	43%	9%	84%	270

Note. The monetary payoffs displayed in the first three columns are in USD in Beard & Beil (1994), in cents of USD in Goeree & Holt (2001), in Yens in Beard et al. (2001) and in Euros in our treatments.

on (from 83% in their treatment 1 to 100% in treatments 3 and 4, for instance), most player As decide not to rely on their partners. The comparison across treatments shows that behaviors highly depends on the size of the stakes: the larger the cost of being unreliant, the higher the reliance rate from player As; the higher the cost of being unreliable, the higher the rate of reliability from player Bs. Beard, Beil, and Mataga (2001) replicate some of these treatments using Japanese subjects. They observe a similar pattern, except that the share of unreliable player Bs (choosing l when they are relied on) is much higher, particularly in treatment 2. Because they implement the (original) sequential form of the game, Beard and Beil (1994) elicit player Bs' behavior only conditional on reliance from player As. It is thus impossible to observe what player Bs matched with unreliant player As would do if their partner acted differently. In contrast, Goeree and Holt (2001) implement the strategic form of the game, and confirm the robustness of previous evidence to this change in the protocol.

In Jacquemet and Zylbersztejn (2010), we assess whether the extent of information about the interaction partner helps to overcome this puzzle. To that end, the experiment implements the normal form of the game – thus eliciting player B's decision unconditional on what player A does – under three information enhancing treatments: simple (perfect stranger) repetition of the game, cheap talk communication about future decisions from player B to player A and observation by player A of the entire history of past decisions taken by player B. The (between subjects) treatments confirm a significant effect of information on coordination on the efficient outcome. At the individual level, the behavior of player Bs appear insensitive to any of the experimental treatments. It means that neither repetition-based learning, nor forward-looking information through messages nor even backward-looking information from observation of past

Table 2: Payoff structure of the one-shot games

	Player B	
Player A	l	r
L	(9.75;3)	(9.75;3)
R	(3;4.75)	(10;5)

(a) Baseline treatment

	Player B	
Player A	l	r
L	(9.75;5)	(9.75;5)
R	(5;9.75)	(10;10)

(b) Egalitarian treatment

decisions manage to discipline player Bs' choices of weakly dominated actions.

The purpose of this short paper is to test whether such puzzling behavior from player Bs is related to the payoff structure of the game.[2] Indeed, in most experimental implementations of this game, the payoffs from the subgame perfect equilibrium is much higher for player A than for player B. Although this does not make B's unreliant decision l a rational answer to A's reliance, non-standard preferences involving aversion to inequality (Fehr and Schmidt, 1999) might be the reason why player Bs forgo efficiency at a personal (and small) monetary cost.

3 Experimental design

We rely on the experimental design of Jacquemet and Zylbersztejn (2010). The Baseline treatment implements the payoff structure used as Treatment 1 in Beard and Beil (1994), presented in Table 2a. We introduce two important changes to Beard and Beil's genuine design. First, we study the effect of learning by repeating the one-shot game 10 times. Each occurrence is one-shot in the sense that: roles are fixed; pairs are rematched in each round using a perfect stranger, round-robin procedure; we avoid the end-game effect by providing no information about the exact number of repetitions; take-home earnings are derived from one round, randomly drawn out of the ten at the end of each experimental session. Second, we elicit both players' decisions in each occurrence of the game. To that matter, we break the original sequentiality of the game and ask each player for unconditional choices in each round. Players are only informed about their own payoffs at the end of each round. All stakes are expressed in Euros, the show-up fee is 5 Euros.

In the Egalitarian treatment, we hold constant all experimental procedures and only change

[2] This hypothesis has been already raised in the literature – see for instance (Goeree and Holt, 2001, p.1412) – but to the best of our knowledge it has never been empirically looked at. One exception is treatment 6 in Beard and Beil (1994), discussed in Section 3. Surprisingly, this treatment is not commented on in the original paper, neither it is discussed as a mean to assess the sensitivity of behaviors to more equalized payoffs. In any case, as stressed above, the original design of Beard and Beil (1994) is inappropriate to study player Bs' behavior since their decisions are elicited only conditional on player A's choice.

the payoff structure of the game towards more symmetry between players. To ease comparison with existing results, we chose the payoff matrix used as Treatment 6 in Beard and Beil (1994), presented in Table 2b. An attractive property of this choice is it modifies the relative payoffs without affecting the main strategic aspects of the game. From player A's point of view, the cost of unreliance (resulting in outcomes (L, l) or (L, r)) is 0.25 Euros as compared to successful coordination (resulting in (R, r)); if player A chooses to rely on player B by selecting R, being reliable (resulting in (R, r)) brings player B a bonus of 0.25 Euros as compared to being unreliable (resulting in (R, l)). The main difference between the two games is the distance between both players' payoffs when the efficient outcome is reached: player B's payoff is much lower in the Baseline Treatment, and exactly the same as player A's gain in the Egalitarian Treatment.

The two treatments are implemented separately, using a between-subject design. For each of the two treatments, we ran three sessions (each involving 20 subjects: 10 As and 10 Bs). Amongst 120 participants, 54 are males and 66 are females. A vast majority of this population (99 subjects) are students with various fields of specialization, 61% of subjects has already taken part in economic experiments run at LEEP. Participants' average age is about 24.[3] Each session lasted about 45 minutes, with an average payoff of 12 Euros in the Baseline Treatment and 17 Euros in the Egalitarian Treatment. No subject participated in more than one experimental session.

Evidence reported by Beard and Beil suggests that under the Egalitarian Treatment, player As tend to be more reliant than in the Baseline (Z=2.70, p<0.01), and player Bs happen to be more reliable when trusted (Z=1.79, p=0.037).[4] From observed behavior in the first round of our treatments (which replicates Beard and Beil setting since subjects are yet unexperienced), we only partly confirm these observations. We do observe that player As are significantly more reliant (Z=1.89, p=0.029), but the likelihood that player Bs happen to be reliable when relied on decreases in our case (Z=1.81, p=0.035); the same result holds once we look at the unconditional probability of being reliable – it falls from 0.8 in the Baseline to 0.63 in the Egalitarian treatment (Z=1.43, p=0.076).

4 Results

Table 3 reports the results from the two treatments, along with a summary of earlier data collected by Beard and Beil (1994). The first three columns provide unconditional decisions observed in each treatment (not available for player B in Beard and Beil, 1994) as well as the rate of reliability conditional on reliance from player A.

Within treatment comparisons provide evidence on the effect of repetition-based learning. We

[3] All sessions took place in the Laboratoire d'Economie Experimentale de Paris (LEEP) at University Paris 1 Panthéon-Sorbonne. The recruitment of subjects makes use of an on-line registration interface adapted from ORSEE (Greiner, 2004) and the experiment is computerized through a software developed under REGATE (Zeiliger, 2000).

[4] Z-statistics and corresponding p-values come from one-tailed tests for equality of proportions.

Table 3: Observed decisions

Game	N	Decisions			Coordination		Failure	
		$Pr(R)$	$Pr(r)$	$P(r\|R)$	(R,r)	(L,l)	(L,r)	(R,l)
Beard and Beil (1994), Tr. 1	35	0.343	–	0.833	0.286	—	—	0.057
Baseline, Round 1	30	0.233	0.800	1.000	0.233	0.200	0.567	0.000
Baseline, Rounds 2-4	90	0.456	0.844	0.902	0.411	0.111	0.433	0.044
Baseline, Rounds 5-7	90	0.589	0.767	0.830	0.489	0.133	0.278	0.100
Baseline, Rounds 8-10	90	0.511	0.811	0.783	0.400	0.078	0.411	0.111
Baseline, overall	300	0.490	0.807	0.844	0.413	0.117	0.393	0.077
Beard and Beil (1994), Tr. 6	26	0.692	–	1.000	0.692	—	—	0.000
Egalitarian, Round 1	30	0.467	0.633	0.643	0.300	0.200	0.333	0.167
Egalitarian, Rounds 2-4	90	0.478	0.733	0.721	0.344	0.134	0.389	0.133
Egalitarian, Rounds 5-7	90	0.456	0.689	0.683	0.311	0.167	0.378	0.144
Egalitarian, Rounds 8-10	90	0.433	0.789	0.795	0.344	0.123	0.444	0.089
Egalitarian, overall	300	0.427	0.727	0.723	0.330	0.147	0.400	0.127

Note. The first three columns provide unconditional decisions observed in each treatment, as well as the rate of reliability conditional on reliance from player A (which is only relevant for Beard and Beil's data). In the last four columns, we present the empirical frequencies of four kinds of possible outcomes (two of which – (L,l) and (L,r) – are not observable in Beard and Beil's data).

observe some learning in both treatments, although its patterns differ substantially from one treatment to the other.[5] In the Baseline Treatment, the differences in player As' behavior between the initial round of the game and rounds 2-4, 5-7, 8-10 are positive and statistically significant (see Model 1 in Table 4: all p<0.01). Moreover, the average effect of repeated interaction on the likelihood of player As' reliant choices is equal to 0.285 (Model 2: p<0.01). Player Bs' preferences, in turn, happen to be highly insensitive to repetition (Model 2: p=0.951). In the Egalitarian Treatment, on the contrary, player As' behavior remains fairly stable over time.[6] The most salient feature in terms of player Bs' behavior is a fairly low rate of reliance in the initial stage of this treatment, which then increases as the game is being repeated; the average effect of repetition on the likelihood of decision r equals 0.104 (Model 2: p=0.064). This increase is however significant only in the first few repetitions of the game (Model 1: p<0.01 for rounds 2-4, p=0.314 and p=0.114 for 5-7 and 8-10).

[5]Due to the structure of our data, the statistical approach must account for correlated observations both within subjects (from one repetition to the other) and across pairs (due to the matching scheme). To address this issue, we rely on OLS clustered-data regressions. We moreover use a delete-one jackknife error correction to avoid small sample biases. The procedure is described in more detail in Jacquemet and Zylbersztejn (2010). The results for each outcome of interest is presented in Table 4.

[6]Model 1 shows that the differences between round 1 and rounds 2-4, 5-7 and 8-10 are highly insignificant, with all p-values around 0.9. Model 2 shows that the average effect of repetition on player As' reliance is insignificant, with p=0.932

Table 4: Parametric tests for equality of proportions.

	Pr(R)		Pr(r)		Pr(R,r)		Pr[$(R,r)\cup(L,l)$]		Pr(L,r)		Pr(R,l)	
Variable	coef	p-value	coef	p-value	coef	p-value	coef	p-value	coef	p-value	coef	p-value
Model 1												
Intercept	0.233	0.002	0.800	0.001	0.233	0.002	0.433	0.007	0.567	0.002	0.000	0.967
ET	0.233	0.007	-0.167	0.270	0.067	0.134	0.067	0.596	-0.233	0.118	0.167	0.007
BT_rounds2-4	0.222	0.000	0.044	0.818	0.178	0.003	0.089	0.618	-0.133	0.483	0.044	0.016
BT_rounds5-7	0.356	0.000	-0.033	0.714	0.256	0.000	0.189	0.183	-0.289	0.044	0.100	0.006
BT_rounds8-10	0.278	0.004	0.011	0.889	0.167	0.033	0.044	0.761	-0.156	0.272	0.111	0.056
ET_rounds2-4	0.011	0.921	0.100	0.006	0.044	0.631	-0.022	0.832	0.056	0.592	-0.033	0.584
ET_rounds5-7	-0.011	0.889	0.056	0.314	0.011	0.929	-0.022	0.859	0.044	0.549	-0.022	0.699
ET_rounds8-10	-0.033	0.872	0.156	0.114	0.044	0.731	-0.033	0.412	0.111	0.391	-0.078	0.479
Model 2												
Intercept	0.233	0.002	0.800	0.001	0.233	0.002	0.433	0.007	0.567	0.002	0.000	0.944
ET	0.233	0.007	-0.167	0.270	0.067	0.134	0.067	0.596	-0.233	0.118	0.167	0.007
BT_rounds2-10	0.285	0.000	0.007	0.951	0.200	0.001	0.107	0.457	-0.193	0.212	0.085	0.001
ET_rounds2-10	-0.011	0.932	0.104	0.064	0.033	0.771	-0.026	0.773	0.070	0.471	-0.044	0.483
Model 3												
Intercept	0.490	0.000	0.807	0.000	0.413	0.001	0.530	0.000	0.393	0.000	0.077	0.001
ET	-0.033	0.731	-0.080	0.533	-0.083	0.493	-0.053	0.260	0.003	0.953	0.050	0.233

Note. Each column summarizes the results of session-clustered (6 clusters in total, 100 observations per cluster, standard errors corrected with a delete-one jackknife) OLS regressions on the outcome described in the first row: player A's decision R, player B's decision r, cooperative outcome (R, r), coordinated outcomes $(R, r) \cup (L, l)$, and decision-mismatches, (L, r) and (R, l). The intercept represents the reference frequency in the Baseline treatment, while dummy ET corresponds to the change in the intercept due to the Egalitarian treatment. All other coefficients are interpreted as the absolute change in the frequency of the dependent variable with respect to the reference point. In Models 1 and 2, explanatory variables are dummies related to the stage of the game (prefix BT stands for the Baseline treatment, ET for the Egalitarian treatment), and the reference is round 1 in the Baseline. Model 3 measures the average effect of the Egalitarian treatment on outcomes as compared to the Baseline.

The inter-treatment comparison confirms that the first period of the game drives most of the observed differences. The only significant difference between the two treatments in terms of player Bs' behavior occurs at the initial stage: the patterns of decision making become very alike in rounds 2-4 (p=0.416), 5-7 (p=0.594), and 8-10 (p=0.888).[7] Finally, aggregate data from rounds 1-10 suggest that all the proportions – related to both players' decisions, as well as resulting outcomes – are statistically the same for both treatments (as shown in Model 3).

We conclude the analysis with a closer look at individual behavior. In Table 5, we group player Bs according to the number of times they choose the weakly dominated decision l over the 10 repetitions of the game. Among the 30 subjects we observe in each treatment, around one third (12 in the Baseline, 11 in the Egalitarian Treatment) always choose the payoff maximizing decision. Among the other two thirds, the number of times the weakly dominated decision is selected is quite dispersed. It happens only once for 4 subjects in both treatments, and otherwise ranges from 2 times to 7 times. In the Egalitarian Treatment, 2 subjects never deviated from the weakly dominated action. This distribution rules out the possibility that the observed failure to maximize payoffs is only a matter of a few individual outliers.

[7] These results come from tests for equality of coefficients in Model 1. For instance, a between-treatment comparison of proportions of decisions r in rounds 2-4 boils down to testing H_0: BT_rounds2-4=ET+CT_rounds2-4.

Table 5: Distribution of player Bs according to the weakly dominated decision

$Pr(l)$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Baseline Treatment	12	4	5	1	3	2	2	1	0	0	0
Egalitarian Treatment	11	4	2	1	5	2	0	3	0	0	2

Note. For each treatment (*in row*) the Table reports the distribution of player Bs according to the number of times the weakly dominated decision l is chosen over the 10 repetitions of the game.

5 Discussion

Accumulated evidence on the experimental game introduced by Beard and Beil (1994) show that a high proportion of subjects fail to select the decision maximizing everyone's monetary payoffs – resulting in an accordingly high share of partners who appear reluctant to rely on others' ability to do so. The main results from this literature are that this behavior is sensitive to the opportunity cost of failure to maximize payoff (Beard and Beil, 1994; Beard, Beil, and Mataga, 2001; Goeree and Holt, 2001), but, as long as the stakes remain small, it is insensitive to the strategic form of the game (normal or sequential, Goeree and Holt, 2001), to repetition-based learning as well to the flows of soft (cheap talk communication) or hard (observation of the decision history) information between players (Jacquemet and Zylbersztejn, 2010). One common property of most experimental implementations of this game is the strong inequality in payoffs between players. This well documented failure to maximize payoff could thus be driven by inequality aversion (as raised, *e.g.*, by Goeree and Holt, 2001).

This short paper reports on an experimental test of this hypothesis. We study a new experimental treatment added to the design of Jacquemet and Zylbersztejn (2010) which neutralizes the effect of relative payoffs comparisons. Despite small differences in observed behavior, this new structure essentially leaves unchanged the outcomes – in terms of both individual decisions and the result of interaction. This non-result thus leaves open the question raised in this paper: why do so many subjects fail to maximize their own payoff through a cognitively costless decision? One remaining possibility is this behavior be one more specificity of WEIRD people, as defined by Henrich, Heine, and Norenzayan (2010). A cross-cultural design based on this experiment would be able to address this hypothesis. We however doubt this could be enough to close the discussion, given in particular that Beard, Beil, and Mataga (2001) explores cultural differences with Japanese subjects. Another possibility is a lack of commitment from subjects towards the experiment they are involved in (see, *e.g.*, Jacquemet, Joule, Luchini, and Shogren, 2009), in the sense that they fail to take the decision problem seriously because the stakes are too low to engage them into the game. This question is next on our agenda.

References

BEARD, T. R., AND J. BEIL, RICHARD O. (1994): "Do People Rely on the Self-Interested Maximization of Others? An Experimental Test," *Management Science*, 40(2), 252–262.

BEARD, T. R., R. O. J. BEIL, AND Y. MATAGA (2001): "Reliant behavior in the United States and Japan," *Economic Inquiry*, 39(2), 270–279.

FEHR, E., AND K. M. SCHMIDT (1999): "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, 114(3), 817–868.

GOEREE, J. K., AND C. A. HOLT (2001): "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions," *American Economic Review*, 91(5), 1402–1422.

GREINER, B. (2004): "An Online Recruitment System for Economic Experiments.," in *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63*, ed. by K. Kremer, and V. Macho, pp. 79–93. Ges. für Wiss. Datenverarbeitung, Göttingen.

HENRICH, J., S. J. HEINE, AND A. NORENZAYAN (2010): "The Weirdest People in the World?," *Behavioral and Brain Sciences*, 33(2/3), 61–83.

JACQUEMET, N., R.-V. JOULE, S. LUCHINI, AND J. SHOGREN (2009): "Earned Wealth, Engaged Bidders? Evidence from a second price auction," *Economics Letters*, 105(1), 36–38.

JACQUEMET, N., AND A. ZYLBERSZTEJN (2010): "Learning, words and actions: experimental evidence on coordination-improving information," *CES Working Paper*, 2010(64).

ROSENTHAL, R. W. (1981): "Games of perfect information, predatory pricing and the chain-store paradox," *Journal of Economic Theory*, 25(1), 92–100.

ZEILIGER, R. (2000): "A presentation of Regate, Internet based Software for Experimental Economics," *http://www.gate.cnrs.fr/ zeiliger/regate/RegateIntro.ppt, GATE.*